Linear Estimation with the unknown scale of noise

$y = Ax + v \qquad Var(v) = \sigma^2 S$

σ^2 is unknown.

$(y, A, \sigma^2 S)$

$$\hat{x} = \left(A^*(\sigma^2 S)^{-1} A\right)^{-1} A^*(\sigma^2 S)^{-1} y$$

$$= \sigma^2 \sigma^{-2} (\quad)^{-1} \cdots$$

$$= (A^* S^{-1} A)^{-1} A^* S^{-1} y$$ Does not depend on σ^2

$$Var(\hat{x}) = \left(A^*(\sigma^2 S)^{-1} A\right)^{-1} = \sigma^2 (A^* S^{-1} A)^{-1}$$

Need to estimate σ^2 : $\hat{\sigma}^2$

$$Q(x) = \|\bar{y} - \bar{A} x\|^2$$

$\bar{y} = \bar{A}x + \bar{v} \qquad \bar{y} = S^{-1/2} y \qquad \bar{A} = S^{-1/2} A$

$\bar{v} = S^{-1/2} v \qquad \bar{v} \sim (0, \sigma^2 I)$

$$Q(x) = \|y - Ax\|^2$$

$$\mathbb{E}\, Q(\hat{x}) \sim \sigma^2$$

$$y = Ax + v \qquad \operatorname{Var} v = \sigma^2 S$$

$$\times\ S^{-1/2}$$

$$\bar{y} = \bar{A}x + \bar{v}$$

$$\bar{y} = S^{-1/2} y \quad \bar{A} = S^{-1/2} A \quad \bar{v} = S^{-1/2} v$$

$$\operatorname{Var} \bar{v} = \sigma^2 I$$

Turn off Bars.

$$y = Ax + v \qquad \operatorname{Var} v = \sigma^2 I$$

$$Q(\hat{x}) = \|y - A\hat{x}\|^2 = \qquad \hat{x} = \underbrace{(A^*A)^{-1}}_{=T} \underbrace{A^*y}_{=z}$$

$$= \|y\|^2 - 2\langle y, A\hat{x}\rangle + \langle A\hat{x}, A\hat{x}\rangle$$

$$= \|y\|^2 - 2\langle \underbrace{A^*y}_{z}, T^{-1}z\rangle + \langle \underbrace{A^*A}_{=T} T^{-1}z, T^{-1}z\rangle \qquad T = A^*A \qquad z = A^*y$$

$$= \|y\|^2 - 2\langle z, T^{-1}z\rangle + \langle z, T^{-1}z\rangle$$

$$= \|y\|^2 - \langle T^{-1}z, z\rangle$$

$$= \|y\|^2 - \langle (A^*A)^{-1}A^*y, A^*y\rangle$$

$$y - A\hat{x} = y - A(A^*A)^{-1}A^*y$$

$$= (I - A(A^*A)^{-1}A^*)\, y \qquad y = Ax + v$$

$$= (I - A(A^*A)^{-1}A^*)(Ax + v)$$

$$= A(I - \underbrace{(A^*A)^{-1}A^*A}_{=I})x$$

$$+ (I - \underbrace{A(A^*A)^{-1}A^*}_{=P})\, v$$

$$P = A(A^*A)^{-1}A^* : \mathcal{R} \to \mathcal{R}$$

$$y - A\hat{x} = (I - P)\, v$$

$\underline{P = P^*}$

$$P^2 = PP = A\underbrace{(A^*A)^{-1}A^*A}_{=I}(A^*A)^{-1}A^* = P$$

$\Rightarrow \underline{P^2 = P}$

$$A : \mathcal{D} \to \mathcal{R}.$$

$$(A^*A)^{-1}A^* = A^- : \mathcal{R} \to \mathcal{D}$$

Pseudoinverse of A

$P = AA^-$ - Orthogonal projector to $\mathcal{R}(A)$ - range of A

$$\mathcal{R}(A) \subset \mathcal{R}$$

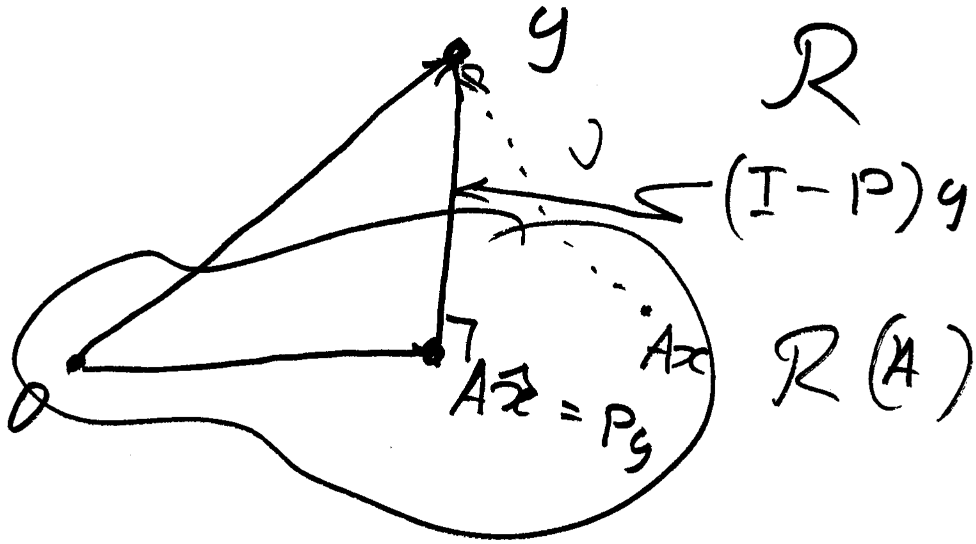
$$E\|y - A\hat{x}\|^2 = E\|(I-P)\nu\|^2$$

$$= tr\left[(I-P)\cdot\sigma^2 I\cdot(I-P)^*\right]$$

$$= \sigma^2 tr\left[(I-P)(I-P)\right] = \sigma^2 tr(I-P)$$

$$I - P - P + \underbrace{P^2}_{=P}$$

$$E\|y - A\hat{x}\|^2 = \sigma^2\cdot tr(I-P)$$



$$\mathcal{R}(A) \subset \mathcal{R}$$

$$A\hat{x} = \underbrace{A(A^*A)^{-1}A^*}_{P}y = Py$$

$*$ if $y \in \mathcal{R}(A) \overset{?}{\Rightarrow} Py = y$

$y \in \mathcal{R}(A) \Leftrightarrow \exists z \in \mathcal{D} : Az = y$

$$Py = PAz = A\underbrace{(A^*A)^{-1}A^*A}_{=I}z = Az = y$$

* if $y \perp \mathcal{R}(A) \overset{?}{\Rightarrow} Py = 0$

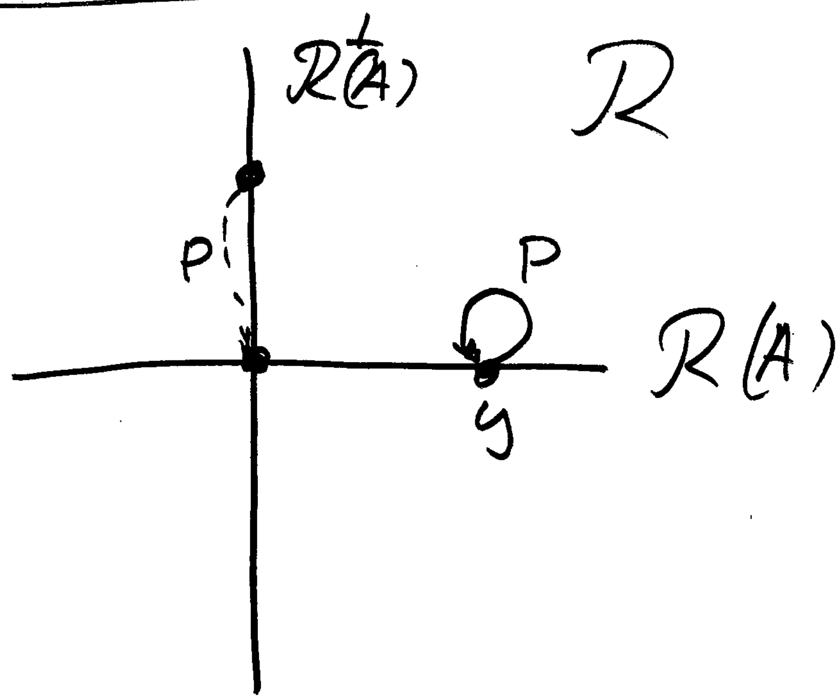
$$y \perp \mathcal{R}(A) \iff \forall z \in \mathcal{R}(A) \ \langle y, z \rangle = 0$$

$$\iff \forall x \in \mathcal{D} \ \langle y, Ax \rangle = 0$$

$$\|Py\|^2 = \langle Py, Py \rangle = \langle Py, y \rangle$$

$$= \langle A \underbrace{(A^*A)^{-1} A^* y}_{= x \in \mathcal{D}}, y \rangle = \langle Ax, y \rangle = 0$$

$$\Rightarrow Py = 0$$

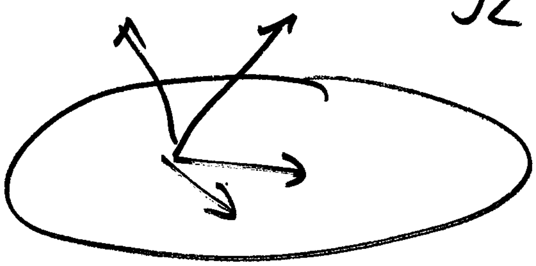


$$\operatorname{tr}(I-P) = \operatorname{tr} I_{\mathcal{R}} - \operatorname{tr} P$$

$$\operatorname{tr} I_{\mathcal{R}} = \operatorname{tr}\begin{bmatrix}1&&&\\&1&&\\&&\ddots&\\&&&1\end{bmatrix} = \sum_{i=1}^{n} 1 = n \qquad n = \dim \mathcal{R}$$

Choose an ONB $\{e_i\}$ $i = 1 .. n$.

$\mathcal{R}$



$\mathcal{R}(A)$

$$\underbrace{e_1 \cdots e_m}_{\mathcal{R}(A)}, \underbrace{e_{m+1} \cdots e_n}_{\mathcal{R}^{\perp}(A)}$$

$m = \dim \mathcal{R}(A)$ $\qquad$ A^*A - invertible.

rank $A = \dim \mathcal{D}$

$\cancel{\mathcal{D}}$ $\dim \mathcal{R}(A) = \underline{\dim \mathcal{D} = m}$

$$\operatorname{tr}\left.\begin{bmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&&0&\\&&&&\ddots\\&&&&&0\end{bmatrix}\begin{matrix}\}\, m\\ \\ \\ \\ \\ \end{matrix}\right\} n \qquad \operatorname{tr} P = m = \dim \mathcal{D}$$

$$Pe_i = \begin{cases} e_i & i \le m \\ 0 & i > m \end{cases}$$

$$E\, Q_{min} = E\|y - A\hat{x}\|^2 = \sigma^2(n-m)$$

$$\hat{\sigma}^2 = \frac{\|y - A\hat{x}\|^2}{n-m}$$

Restore Bars

$$\hat{\sigma}^2 = \frac{\|\bar{y} - \bar{A}\hat{x}\|^2}{n-m} \qquad \bar{y} = S^{-1/2} y \qquad \bar{A} = S^{-1/2} A$$

$$\|\bar{y} - \bar{A}\hat{x}\|^2 =$$

$$= \underbrace{\|S^{-1/2}\bar{y}\|^2}_{=\langle S^{-1/2}y,\, S^{-1/2}y\rangle} - 2\langle S^{-1/2}y,\, S^{-1/2}A\hat{x}\rangle - \|S^{-1/2}A\hat{x}\|^2$$

$$= \langle S^{-1}y, y\rangle - 2\langle \underbrace{A^*S^{-1}y}_{=z},\, \underbrace{(A^*S^{-1}A)^{-1}}_{=T}\underbrace{A^*S^{-1}y}_{z}\rangle + \langle \underbrace{A^*S^{-1}A}_{=T}\hat{x}, \hat{x}\rangle$$

$$= \langle S^{-1}y, y\rangle - 2\langle z, T^{-1}z\rangle + \langle \underbrace{TT^{-1}}_{=I}z, T^{-1}z\rangle$$

$$= \underbrace{\langle S^{-1}y, y\rangle}_{=w} - \langle T^{-1}z, z\rangle$$

$$= w - \langle Tz, z\rangle$$

$$\|\bar{y} - \bar{A}\hat{x}\| = w - \langle T^{-1}z, z\rangle$$

$$w = \langle S^{-1}y, y\rangle$$

$$T = A^* S^{-1} A$$

$$z = A^* S^{-1} y$$

* $\hat{\sigma}^2 = \dfrac{w - \langle T^{-1}z, z\rangle}{n-m}$

* $\hat{x} = T^{-1}z$

* $\widehat{\mathrm{var}}(\hat{x}) = \hat{\sigma}^2 T^{-1} = \dfrac{w - \langle T^{-1}z, z\rangle}{n-m} T^{-1}$

$$(y, A, \underset{\sim}{\sigma^2} S) \longmapsto (T, z, w, n)$$

$$\begin{matrix} (y_1, A_1, \underset{\sim}{\sigma^2} S_1) \\ (y_2, A_2, \underset{\sim}{\sigma^2} S_2) \end{matrix} \oplus \left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \underset{\sim}{\sigma^2} \underbrace{\begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix}}_{S} \right)$$

$$\begin{matrix} (T_1, z_1, w_1, n_1) \\ (T_2, z, w_2, n_2) \end{matrix} \oplus \to (T, z, w, n) = ?$$

$$T = T_1 + T_2 \quad , \quad z = z_1 + z_2$$

$$n_1 = \dim y_1 \qquad n_2 = \dim y_2$$

$$n = \dim \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = n_1 + n_2$$

$$w = \langle S^{-1} y, y \rangle = \left\langle \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\rangle$$

$$= \left\langle \underbrace{\begin{bmatrix} S_1^{-1} & 0 \\ 0 & S_2^{-1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}}_{\begin{bmatrix} S_1^{-1} y_1 \\ S_2^{-1} y_2 \end{bmatrix}}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\rangle$$

$$= \langle S_1^{-1} y_1, y_1 \rangle + \langle S_2^{-1} y_2, y_2 \rangle$$

$$= w_1 + w_2$$

$$\langle T_1, z_1, w_1, n_2 \rangle + \langle T_2, z_2, w_2, n_2 \rangle =$$

$$\langle T_1 + T_2, z_1 + z_2, w_1 + w_2, n_1 + n_2 \rangle$$

$$\hat{x} = T^{-1} z$$

$$\widehat{\mathrm{var}}(\hat{x}) = \frac{w - \langle T^{-1} z, z \rangle}{n - m} T^{-1}$$

Problems with uncertainty in A.

$$y = \underset{\sim ?}{A} x + v$$

Simple example.

$$y = a x + v \qquad v \sim (0, S)$$

$$x \sim (x_0, F) \quad \text{a priori info}$$

$$a \sim (a_0, G)$$

x, v, a independent.

$\hat{x} = Ry + r$ – linear est of x

$$E(\hat{x} - x)^2 = H(R, r) \sim \min_{R, r}$$

$$\min_{R,r} = F_1 = \underbrace{\frac{GF + G x_0^2 + S}{GF + G x_0^2 + S + a_0^2 F}}_{= C} F$$

F_1 – a posteriori variance.

$$F_1 = CF \leq F \qquad c < 1$$

(a) a priori info vanishes : $F \to +\infty$

$$C \to \frac{G}{G + a_0^2} \qquad F_1 = CF \to +\infty$$

(C → fin., F → $+\infty$)

(b) $a = a_0$ (known precisely) $\Rightarrow G = 0$

$$C = \frac{S}{S + a_0^2 F} \to 0 \quad F_1 = CF \to \frac{S}{a_0^2} \qquad F \to \infty$$

Uncertainty in A – general case.

$$y = Ax + v \qquad v \sim (0, S)$$

$x \sim (x_0, F)$ a priori info

$E A = A_0$

and something else.

$$\hat{x} = \underline{R}\, y + \underline{r}$$

$$\hat{x} - x = R(Ax + v) + r - x$$

$$\left|\begin{array}{lll} x = x_0 + x' & E x' = 0 & \mathrm{Var}\, x' = F \\ A = A_0 + A' & E A' = 0 & \end{array}\right.$$

$$\hat{x} - x = \left(R(A_0 + A') - I\right)x + r + Rv$$

$$= RA'x + (RA_0 - I)(x_0 + x') + r + Rv$$

$$= RA'x + (RA_0 - I)x' + \underbrace{\left[(RA_0 - I)x_0 + r\right]}_{=0} + Rv$$

take: $r = -(RA_0 - I)x_0$

$$\Rightarrow \hat{x} - x = RA'x + (RA_0 - I)x' + Rv$$

$$E_{A,x,v}\|\hat{x}-x\|^2 = E_{A,x}\|RA'x\|^2 + E_x\|(RA_0-I)x'\|^2 + E_v\|Rv\|^2$$

$$\underbrace{E_{A,x}}_{E_x E_A}\langle RA'x, (RA_0-I)x'\rangle =$$

$$= E_x\langle R\cdot \underbrace{E_A A'}_{=0}\cdot x, (RA_0-I)x'\rangle =$$

$$= E_x\, 0 = 0$$

$$E_v\|Rv\|^2 = tr\, RSR^*$$

$$E_x\|(RA_0-I)x'\|^2 = tr(RA_0-I)F(RA_0-I)^*$$

$$E_{A,x}\|RA'x\|^2 = ?$$

$$E_x\|\underbrace{RA'x}_{z}\|^2 = E_x\, tr\, zz^*$$

$$\|z\|^2 = tr\, zz^* = tr\begin{bmatrix} z_1 \\ \vdots \\ z_m\end{bmatrix}[z_1 \ldots z_m]$$

$$= tr\begin{bmatrix} z_1^2 & & \cdots \\ \vdots & z_2^2 & \ddots \\ \cdots & & z_m^2\end{bmatrix} = \sum_{i=1}^{m} z_i^2$$

$$z = R A' x$$

$$E_x \| z \|^2 = E_x \operatorname{tr} R A' \underbrace{x \cdot x^*} A'^* R^*$$

$$= \operatorname{tr} R A' \left[E_x x x^* \right] A'^* R^*$$

$E_x x x^*$ – second moment of x

$$E_x x x^* = E (x_0 + x')(x_0 + x')^*$$

$$= x_0 x_0^* + E x_0 \underline{\underline{x'^*}} + \underbrace{E x' x_0^*}_{=0} + \underbrace{E x' x'^*}_{=F}$$

$$= x_0 x_0^* + F = \bar{F}$$

$$E \begin{bmatrix} x'_1 \\ \vdots \\ x'_m \end{bmatrix} \begin{bmatrix} x'_1 & \cdots & x'_m \end{bmatrix} = E \begin{bmatrix} x'^2_1 & x'_1 x'_2 & \cdots & x'_1 x'_m \\ \vdots & & & \\ x'_m x'_1 & \cdots & \cdots & x'^2_m \end{bmatrix}$$

$$= \operatorname{Var} x = F$$

$$\underbrace{E_x \| R A' x \|^2} = \operatorname{tr} R \underbrace{A' \bar{F} A'^*} R^*$$

$$E_A(\checkmark) = \operatorname{tr} R \cdot \underbrace{E A' \bar{F} A'^*}_{=J} \cdot R^* = \operatorname{tr} R J R^*$$

$$J = E_A \left(A' \bar{F} A'^* \right)$$

$$E\|\hat{x}-x\|^2 = \mathrm{tr}\left[RJR^* + (RA_0 - I)F(RA_0 - I)^* + RSR^*\right] = \mathrm{tr}\,Q$$

$$Q = (RA_0 - I)F(RA_0 - I)^* + R(S+J)R^*$$

$$\mathrm{tr}\,Q \sim \min_R$$

For problem w. a priori info

$$Q = (R\underline{A} - I)F(R\underline{A} - I)^* + R\underline{S}R^*$$

$$A \to A_0 \qquad S \to S + J$$

$Q \sim \min\limits_R$:

$$Q = \left(A_0^*(S+J)^{-1}A_0 + F^{-1}\right)^{-1}$$

$$R = Q A_0^*(S+J)^{-1}$$

$$r = Q F^{-1} x_0$$

$$\hat{x} = Q\left(A_0^*(S+J)^{-1}y + F^{-1}x_0\right)$$

$$\mathrm{Var}(\hat{x} - x) = Q$$

$$\mathrm{Var}(\hat{x}_i - x_i) = Q_{ii}$$

Examples

$$y = Ax + v \qquad v \sim (0, \sigma^2 I)$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$x_0 = \begin{bmatrix} c \\ c \end{bmatrix} \quad F = \varphi^2 I$$

(a) $$y_1 = (1+\varepsilon_1)x_1 + v_1 \qquad \varepsilon_i \sim (0, \delta^2)$$
$$y_2 = (1+\varepsilon_2)x_2 + v_2 \qquad \text{indep.}$$

$$A = \begin{bmatrix} 1+\varepsilon_1 & 0 \\ 0 & 1+\varepsilon_2 \end{bmatrix} = I + \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} = A_0 + A'$$

$$J = E\, A'(F + x_0 x_0^T) A'^T$$

$$= E \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} \cdot \left(\varphi^2 I + \begin{bmatrix} c \\ c \end{bmatrix} [c \ \ c] \right) \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix}$$

$$= \varphi^2 E \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} + E c^2 \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix}$$

$$= \varphi^2 E \begin{bmatrix} \varepsilon_1^2 & 0 \\ 0 & \varepsilon_2^2 \end{bmatrix} + c^2 E \begin{bmatrix} \varepsilon_1 & \varepsilon_1 \\ \varepsilon_2 & \varepsilon_2 \end{bmatrix} \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix}$$

$$= \varphi^2 \delta^2 I + c^2 E \begin{bmatrix} \varepsilon_1^2 & \varepsilon_1 \varepsilon_2 \\ \varepsilon_2 \varepsilon_1 & \varepsilon_2^2 \end{bmatrix} = \varphi^2 \delta^2 I + c^2 \delta^2 I$$